Chapter 14: Hidden Variables

May 26th, 2018

Eliana Duarte

OVGU

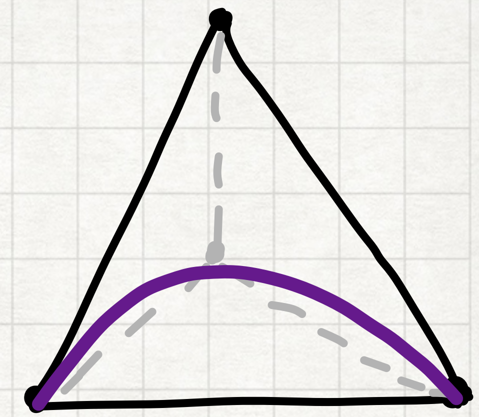
- A random variable in a model is hidden or latent if it is not measured.

- A random variable in a model is hidden or latent if it is not measured.
- In a hidden variable model this means that the probability densities on the observed random variables are obtained by computing marginals of the joint distribution of a fully observed model.

- A random variable in a model is hidden or latent if it is not measured.
- In a hidden variable model this means that the probability densities on the observed random variables are obtained by computing marginals of the joint distribution of a fully observed model.
- Hidden variable models are usually more complex:
  - → semialgebraic description (Example. 14.1.7)
  - → singularities (Proposition 14.1.8)

14.1. MIXTURE MODELS

Example 3.2.9: Binomial Random variable



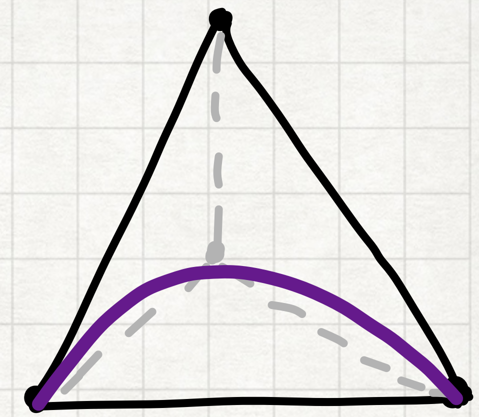
- Consider a biased coin, $P(H) = \theta$, $P(T) = 1 - \theta$

  X = "number of heads in two trials"

  $\mathcal{B}_2 = \{ ((1-\theta)^2, 2(1-\theta)\theta, \theta^2) : 0 \leq \theta \leq 1 \}$ → Parametric Description

  $= \{ (p_0, p_1, p_2) \in \mathbb{R}^3 : p_0 + p_1 + p_2 = 1,\ p_2^2 - 4p_1p_3 = 0,\ p_0, p_1, p_2 \geq 0 \}$ → Implicit Description.

Example 3.2.9: Binomial Random variable



- Consider a biased coin, $P(H)=\theta$, $P(T)=1-\theta$

  $X=$ "number of heads in two trials"

  $\mathcal{B}_2 = \{ ((1-\theta)^2, 2(1-\theta)\theta, \theta^2) : 0 \leq \theta \leq 1 \} \rightarrow$ Parametric Description

  $= \{ (p_0, p_1, p_2) \in \mathbb{R}^3 : p_0+p_1+p_2=1,\ p_2^2-4p_1p_3=0,\ p_0, p_1, p_2 \geq 0 \} \rightarrow$ Implicit Description.

- Suppose we have two coins C_1, C_2, we select one at random, toss the coin twice and record # of heads, Y

$$P(Y=0) = \underbrace{P(C=C_1)}_{\lambda}\underbrace{P(X_{C_1}=0)}_{(1-\theta_1)^2} + \underbrace{P(C=C_2)}_{(1-\lambda)}\underbrace{P(X_{C_2}=0)}_{(1-\theta_2)^2}$$

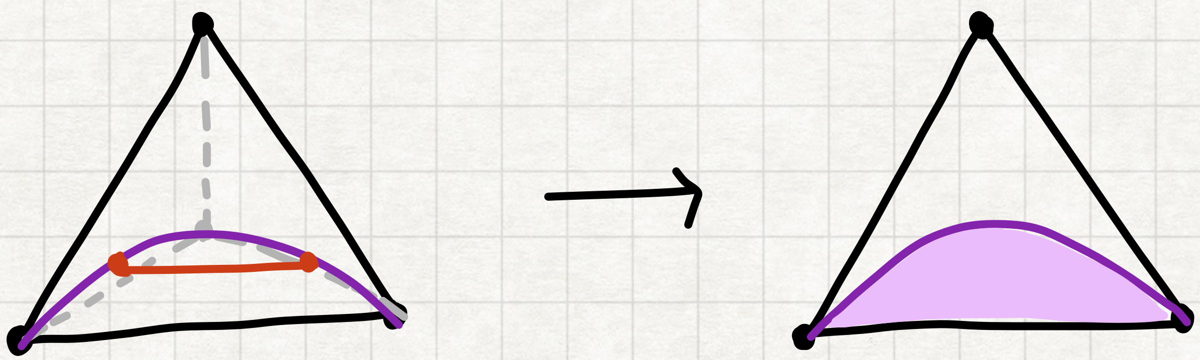
$$= \lambda\cdot(1-\theta_1)^2 + (1-\lambda)(1-\theta_2)^2$$

- The distribution of Y is given by

$$\begin{pmatrix} p_0 \\ p_1 \\ p_2 \end{pmatrix} = \lambda \begin{pmatrix} (1-\theta_1)^2 \\ 2\cdot(1-\theta_1)\,\theta_1 \\ \theta_1^2 \end{pmatrix} + (1-\lambda) \begin{pmatrix} (1-\theta_2)^2 \\ 2\cdot(1-\theta_2)\,\theta_2 \\ \theta_2^2 \end{pmatrix}, \qquad \begin{matrix} 0 \le \lambda \le 1 \\ 0 \le \theta_1 \le 1 \\ 0 \le \theta_2 \le 1 \end{matrix}$$

$\downarrow$

Parametric Description



$$\Delta_2 \cap \left\{ (p_0, p_1, p_2) : \det\begin{pmatrix} 2p_0 & p_1 \\ p_1 & 2p_2 \end{pmatrix} \geq 0 \right\} \rightarrow$$

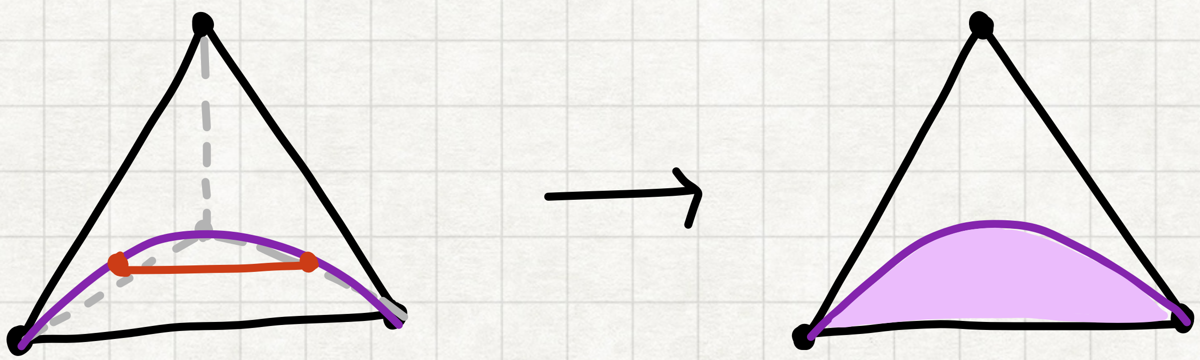
Implicit semialgebraic description

- The distribution of Y is given by

$$\begin{pmatrix} p_0 \\ p_1 \\ p_2 \end{pmatrix} = \lambda \begin{pmatrix} (1-\theta_1)^2 \\ 2\cdot(1-\theta_1)\,\theta_1 \\ \theta_1^2 \end{pmatrix} + (1-\lambda) \begin{pmatrix} (1-\theta_2)^2 \\ 2\cdot(1-\theta_2)\,\theta_2 \\ \theta_2^2 \end{pmatrix}, \qquad \begin{matrix} 0 \le \lambda \le 1 \\ 0 \le \theta_1 \le 1 \\ 0 \le \theta_2 \le 1 \end{matrix}$$

$\downarrow$

Parametric Description



$$\Delta_2 \cap \left\{ (p_0, p_1, p_2) : \det \begin{pmatrix} 2p_0 & p_1 \\ p_1 & 2p_2 \end{pmatrix} \geq 0 \right\} \rightarrow$$

Implicit semialgebraic description

- This model is denoted by $\mathrm{Mixt}^2(\mathcal{B}_2)$ and is the two mixture model of a binomial random variable with two trials.

- Consider a discrete model $\mathcal{M} \subseteq \Delta_{r-1}$, let X be the r.v. modeled by $\mathcal{M}$.

Definition 14.1.1: The K-th mixture model of $\mathcal{M} \subseteq \Delta_{r-1}$ is the family of probability distributions

$$\mathrm{Mixt}^K(\mathcal{M}) = \{ \pi_1 P^1 + \cdots + \pi_K P^K : \pi \in \Delta_{K-1}, P^1, \ldots, P^K \in \mathcal{M} \}$$

- Consider a discrete model $\mathcal{M} \subseteq \Delta_{r-1}$, let X be the r.v. modeled by $\mathcal{M}$.

Definition 14.1.1: The K-th mixture model of $\mathcal{M} \subseteq \Delta_{r-1}$ is the family of probability distributions

$$\text{Mixt}^K(\mathcal{M}) = \left\{ \pi_1 P^1 + \cdots + \pi_K P^K : \ \pi \in \Delta_{K-1},\ P^1, \dots, P^K \in \mathcal{M} \right\}$$

- Interpretation: → A population is divided into K groups.
  - → The i-th group follows a distribution P^i.
  - → The hidden variable H is $P(H=i) = \pi_i$

- Consider a discrete model $\mathcal{M} \subseteq \Delta_{r-1}$, let X be the r.v. modeled by $\mathcal{M}$.

Definition 14.1.1: The K-th mixture model of $\mathcal{M} \subseteq \Delta_{r-1}$ is the family of probability distributions

$$\mathrm{Mixt}^K(\mathcal{M}) = \{\pi_1 \mathbf{p}^1 + \cdots + \pi_K \mathbf{p}^K : \pi \in \Delta_{K-1}, \mathbf{p}^1, \ldots, \mathbf{p}^K \in \mathcal{M}\}$$

- Interpretation: → A population is divided into K groups.
  → The i-th group follows a distribution $\mathbf{p}^i$.
  → The hidden variable H is $P(H=i) = \pi_i$
- In Example 3.2.9. the hidden variable is which coin we choose.
- In a mixture model, we suppose that for some

$$\pi = (\pi_1, \ldots, \pi_K) \in \Delta_{K-1} \text{ and } \mathbf{p}^1, \ldots, \mathbf{p}^K \in \mathcal{M}$$

$$P(H=i) = \pi_i \quad , \quad P(X=j \mid H=i) = p_j^i$$

$$\Rightarrow P(X=j) = \sum_{i=1}^{K} \pi_i p_j^i$$

Example 14.1.2: Mixture of independence model

Consider two random variables

$\Omega_X = \{\text{Never, sometimes, Frequently}\}$

X = "How much a person watches soccer"

$\Omega_Y = \{\text{Bald, Short, Medium, Long}\}$

Y = "How much hair a person has".

- It is expected that $X \perp\!\!\!\perp Y$, however it is found that people with short hair watch more soccer.
- But within each gender group (men and women) we have $X \perp\!\!\!\perp Y$.
- If we only observe the joint distribution of X and Y, we would observe a distribution in $\mathrm{Mixt}^2(\mathcal{M}_{X \perp\!\!\!\perp Y})$.